



## GENOME WATCH

## Looking at Beijing's skyline

Dorota Jamrozcy and Teemu Kallonen

This month's Genome Watch explores recent findings on the global population structure and the origin of the successful *Mycobacterium tuberculosis* Beijing lineage that have been revealed by high-throughput sequencing of global bacterial collections.

*Mycobacterium tuberculosis* and other members of the *M. tuberculosis* complex (MTBC) are major human pathogens that are the causative agents of human tuberculosis and are responsible for over 1 million deaths worldwide every year<sup>1</sup>. The *M. tuberculosis* Beijing lineage has been of particular concern owing to its contribution to outbreaks of multidrug-resistant tuberculosis (MDR-TB) around the world. Now, two new studies use whole-genome sequencing (WGS) to analyse the origin and global population structure of the *M. tuberculosis* Beijing lineage and propose that this lineage originated in East Asia and that its expansion correlated to changes in the human population<sup>2,3</sup>.

Merker *et al.*<sup>2</sup> analysed 4,987 *M. tuberculosis* Beijing lineage isolates from 99 countries. By applying the mycobacterial interspersed repetitive unit-variable-number tandem repeat (MIRU-VNTR) typing method, they resolved the collection into six major clonal complexes and one basal sub-lineage, consisting of three clusters. The study found that isolates from a region defined as East Asia and the Far East contained the highest diversity of clonal complexes, suggesting that this region represent the most likely origin of the Beijing lineage. To verify the findings, 110 representative isolates were

whole-genome sequenced with the Illumina MiSeq platform for phylogenetic reconstruction. A midpoint-rooted phylogenetic tree revealed three minor Asian ancestral clusters and five modern clades that were broadly defined as Asian Africa 1 and 2, Pacific, Europe-Russia and Central Asia. Assignment of ancestral and modern clusters based on tree topology correlated to the distribution of previously identified regions of difference that define sub-groups within the Beijing lineage<sup>4</sup>. The authors further utilized the WGS data to generate a Bayesian skyline plot of temporal changes in the population size of the Beijing lineage, using a model based on the contemporary mutation rate of *M. tuberculosis*. The plot revealed two rapid population growth events that were proposed to coincide with the Industrial Revolution and with the First World War, and a third, more recent smaller expansion around the time of the HIV epidemic and first outbreaks of MDR-TB in the 1990s.

In a separate study, Luo *et al.*<sup>3</sup> sequenced 95 isolates of MTBC lineage 2, which consists predominantly of Beijing strains, using the Illumina HiSeq platform. Combined with 263 previously sequenced genomes, they analysed 358 globally derived isolates, the majority originating from China and Russia. The resulting phylogenetic tree revealed that the collection diverged early into a minor proto-Beijing group and the dominant Beijing group. Isolates were grouped into one of four regional categories: southern East Asia, northern East Asia, Southern Asia and Northern Asia. Isolates derived from southern East Asia were widely distributed across the phylogeny, leading the authors to share the conclusion of Merker *et al.*<sup>2</sup> that the Beijing lineage originated in East Asia. The authors also conducted a Bayesian skyline analysis of temporal changes in the effective population size of the MTBC lineage 2, but based their analysis on a previously

described model that estimates the origin of MTBC at approximately 70,000 years ago<sup>5</sup>. This led the authors to propose that a marked population size expansion of the MTBC lineage 2 occurred in the Neolithic era during a period of expansion of the human population that resulted from agricultural transition in China, around 6,500 years ago.

These studies illustrate the power of WGS when investigating the origin and spread of an infectious disease and highlight the potential of using next-generation sequencing to detect temporal changes in the effective bacterial population size to identify points in time when populations rapidly expanded or decreased in size. Furthermore, these data strengthen the hypothesis that the Beijing lineage originated in East Asia and that changes in the human population have coincided with increases in the effective population size of this important pathogen. However, as different mutation rates were selected for the Bayesian skyline analyses in the different studies, the effective population increases were correlated to different demographic events, highlighting the importance of model selection when dating bacterial evolution.

Dorota Jamrozcy and Teemu Kallonen are at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

e-mail: [microbes@sanger.ac.uk](mailto:microbes@sanger.ac.uk)

doi:10.1038/nrmicro3536

Published online 10 August 2015

1. World Health Organization. *Global tuberculosis report 2014* (WHO, 2014).
2. Merker, M. *et al.* Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat. Genet.* **47**, 242–249 (2015).
3. Luo, T. *et al.* Southern East Asian origin and coexpansion of *Mycobacterium tuberculosis* Beijing family with Han Chinese. *Proc. Natl Acad. Sci. USA* **112**, 8136–8141 (2015).
4. Tsolaki, A. G. *et al.* Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **43**, 3185–3191 (2005).
5. Comas, I. *et al.* Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat. Genet.* **45**, 1176–1182 (2013).

## Competing interests statement

The author declares no competing interests.

